DOCUMENT RESUME

ED 160 096 IR 006 343

AUTHOR	Shetler, Stanwyn G.
TITLE	The Flora North America Generalized System for Describing the Morphology of Organisms. Research Report No. 4.
INSTITUTION	Margaret Woodbury Strong Museum, Rochester, N.Y. Museum Data Bank Committee.
PUB DATE	Jan 75
NOTE	25p.; Best copy available
EDRS PRICE	MF-$0.83 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Botany; *Data Bases; *Electronic Data Processing; *Information Storage; Numbers; Organization; Records (Forms); Tables (Data); *Taxonomy
IDENTIFIERS	*Flora North America

ABSTRACT

The file organization for the computerized Flora North America (FNA) data bank is described. A four level character hierarchy allows subdivision of any flora description into as many as four levels in order to specify plant character precisely. Terms at any one level will not necessarily be parallel in status. Both PLANTS and LEAVES serve as primary headings, since these are the terms by which the data base is searched or sorted. At least two levels, the first and fourth, must always be specified because a character cannot be expressed without both. The second and third are optional. Any larger hierarchy must be condensed to four levels. The actual values or character-states which describe the characters, are appended to the lowest (fourth), most specific level. A character can be described by any number of character-states which may be expressed in words or numbers. The printing of this computerized record reads like conventional language due to assigning of print and non-print codes, and the use of parallel numerical classification codes for all terms at all levels. This numeriClature process is fully described. Appendices include the standard FNA form for compiling morphological data and a sample record in FNA format. (Author/MBR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED160096

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MUSEUM DATA BANK RESEARCH REPORT

Number 4

The Flora North America Generalized System for Describing the Morphology of Organisms

by

Stanwyn G. Shetler

January 1975

This is an abridged preprint of the paper, "A Generalized Descriptive Data Bank as a Basis for Computer-Assisted Identification," which was presented in a symposium on "Automatic Identification" at Cambridge University in September 1973.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. G. Chenhall

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Museum Data Bank Committee
Robert G. Chenhall, Chairman

Margaret Woodbury Strong Museum
700 Allen Creek Road
Rochester, N.Y.

IR004343

The major principles governing morphological description in the FNA System are set forth by Shetler et al. (1973), and the specific methods are explained by Porter et al. (1973), in the first and second parts, respectively, of A Guide for Contributors to Flora North America. Included in Part II, compiled by Porter et al., is a provisional edition--outlined in accepted, classified order--of the exhaustive lists of morphological characters (28,000) presently contained in the Morphological/Ecological Vocabulary (MEV), an authority file of the FNA Data Bank. Of the 28,000 potential characters, the Guidebook specifies about 350, insofar as they apply, as constituting the minimum number required in every FNA description. On the average, probably no more than 25% of these actually apply in a particular case. The rationale for a minimum requirement of specified characters is that only by this means can true comparability from one description to another be assured.

Characters (descriptors)have a nested relationship to each other. Thus, for example, the stem is only one part of the whole plant, and the leaf in turn attaches to the stem; the rootlet is a part of the root, which in turn is a part of the plant, etc. This character-nesting or hierarchy may be simple, as in describing the duration or habit of the plant as a whole, but the nesting also may be complicated, breaking down into many levels from the overall plant to the ultimate part or character. The levels of nesting that one recognizes in any classified hierarchy of characters must be kept, in the final analysis, to a relatively small, arbitrary number, because, intrinsically, there is no absolute limit. A compromise must be struck between the two-level situation of describing plant habit, for example, and the multi-level situation that can arise in describing, for instance, the vestiture of a particular surface of a particular plant part.

Thus, for practical purposes, the potentially open-ended hierarchy must be closed--telescoped into a few arbitrarily fixed levels. Consequently, the terms of any one level of headings or subdivisions will not necessarily be parallel in status. Both PLANTS and LEAVES, for example, may serve as primary headings. A purely functional point of view must be taken. These are the terms by which the data base is to be searched or sorted or by which the descriptions are to be organized in printed form, as in a Flora.

Evaluation of the nesting of a representative cross-section of taxonomic characters led to the adoption of a four-level character hierarchy. This is to say that a plant may be subdivided into as many as four levels of structure in order to specify a character precisely. At least two levels must be indicated, because a character cannot be expressed without specifying the first and fourth levels, but the intermediate second and third levels are optional. Any larger hierarchy must be condensed to four levels as naturally as possible. The actual values or character-states, which describe the characters, are appended to the lowest (fourth), most specific level.

A given character may be described by any number of character-states, and the character-states may be expressed in words or numbers, including statistics. Any number of modifiers (qualifiers) may be appended to each character-state, provided that approved modifiers are used.

Three examples will serve to illustrate the principle of the four-level character hierarchy:

Example A. Cauline leaves usually whorled, sometimes opposite

Level 1.	CAULINE LEAVES			
Level 2.	--			
Level 3.	--			
Level 4.	ARRANGEMENT	WHORLED	USUALLY	
		OPPOSITE	SOMETIMES	

Example B. Ovaries 3-6

Level 1.	FLOWERS		
Level 2.	GYNOECIUM		
Level 3.	OVARIES		
Level 4.	NUMBER	3-6	

Example C. Outer surface of sepals densely lanate or tomentose

Level 1.	FLOWERS			
Level 2.	SEPALS			
Level 3.	ABAXIAL SURFACES			
Level 4.	VESTITURE TYPE	LANATE	DENSELY	
		TOMENTOSE	DENSELY	

If these character descriptions were to be printed out by computer exactly as stored with the full hierarchy explicitly stated, they would read quite unnaturally:

(A) CAULINE LEAVES ARRANGEMENT WHORLED USUALLY OPPOSITE SOMETIMES

(B) FLOWERS GYNOECIUM OVARIES NUMBER 3-6

(C) FLOWERS SEPALS ABAXIAL SURFACES VESTITURE TYPE LANATE DENSELY TOMENTOSE DENSELY

In fact, the scheme was devised on the assumption that only selected words would be printed. This explains why certain words are plural, others singular. Selective printing was to be accomplished by including a print or not-print code with each level in every character description as stored in the computer.

With selective printing, these character descriptions would read almost like conventional descriptions:

(A) CAULINE LEAVES WHORLED USUALLY OPPOSITE SOMETIMES

(B) OVARIES 3-6

(C) [SEPALS--understood by context] ABAXIAL SURFACES LANATE DENSELY TOMENTOSE DENSELY

Finally, it was expected that the report-generating program would be refined still further until word order could be reversed and necessary

punctuation and connectives could be added as desired. Modifiers then could be placed in their conventional position, and the character-states could be separated by the appropriate punctuation marks ("cauline leaves usually whorled, sometimes opposite" instead of "cauline leaves whorled usually opposite sometimes").

A concrete example of a species description is illustrated in tabular form in Table 1. It is a partial description of Xylosma flexuosa (H.B.K.) Hemsl. (Flacourtiaceae), the "Expanded Description" in Porter et al. (1973, p. viii). The tabular organization approximates the logical structure of a description in the data bank but should not be construed as the physical representation in magnetic storage.

Table 1 provides a vivid example of the order of magnitude of expansion when a typical, telegraphic, taxonomic description is blown up into data bank format. Everything must be made explicit. Thus, while to the botanist duration is implicit in the term "shrub," the computer must be "told" that the plant is a perennial. Otherwise, any later machine search for perennials will pass by this species. Likewise, it is necessary to state explicitly that such plant parts as branches, twigs, and leaves are "present."

The conventional version of this description, as it would be printed in a Flora or monograph, consists of the italicized terms in the table (reading from left to right and top to bottom) and would print out approximately as follows, if one assumes that the report-generating program would correct word order and insert essential punctuation:

> SHRUBS TO 2 M HIGH, POLYGAMODIOECIOUS; WOOD MODERATELY HARD, DIFFUSE-POROUS; BRANCHES SLENDER OR STOUT, GLABRATE; TWIGS DENSELY PUBERULENT OR GLABRESCENT; SPURS OFTEN PRESENT; THORNS OFTEN PRESENT, TO 33 MM LONG, SLENDER OR STOUT; LEAVES REMOTE OR SOMETIMES CROWDED (ON SPURS); BLADES RHOMBIC-OVATE, OVATE-ELLIPTIC, ELLIPTIC, OBLANCEOLATE OR OBOVATE, CHARTACEOUS OR SUBCORIACEOUS, TO 6 CM LONG, TO 3 CM WIDE; APEX ACUMINATE, ACUTE, OBTUSE OR ROUNDED; BASE ATTENUATE-CUNEATE OR CUNEATE; MARGIN IRREGULARLY CRENULATE-SERRULATE; ADAXIAL SURFACES GREEN, OFTEN LUSTROUS, GLABROUS; PRIMARY VEINS PUBERULENT: SECONDARY VEINS 5-6 PER SIDE, EVIDENT; TERTIARY VEINS INCONSPICUOUS; ABAXIAL SURFACES PALER GREEN, GLABROUS; SECONDARY VEINS EVIDENT; TERTIARY VEINS INCONSPICUOUS; PETIOLES SHORT; INFLORESCENCES FASCICULATE, SESSILE; PEDICELS TO 5 MM LONG, GLABRATE OR PUBERULENT;SEPALS 5, OVATE-OBLONG, 1-2 MM LONG; APEX ACUTE OR OBTUSE; MARGIN CONSPICUOUSLY CILIATE; ABAXIAL SURFACES GLABROUS OR PUBERULENT;FILAMENTS 3-4 MM LONG; CARPELS 2-6; OVULE NUMBER PER LOCULE 2-8; STYLES PERSISTENT; STYLE BRANCHES 2; STIGMAS 2, DILATED; FRUITS SUBGLUCOSE, TO 6 MM DIAMETER, RED, GLABROUS; SEEDS 3-4 MM LONG, 2-3 MM DIAMETER; SEED NUMBER PER FRUIT 2-8.

The diagnostic features embedded in this description are the ones starred (*) in Table 1 and of course would compile into a shorter paragraph.

This data format imposes four main requirements on the computer-processing strategy. First, all relationships of the character classification must be maintained, which is to say that both the hierarchy (nesting) and the ordering (arrangement) of the characters must be held constant throughout all taxonomic descriptions in the data bank. Second, every

character must constitute a unique combination of terms in levels 1 to 4 (i.e., primary part, secondary part, tertiary part, character), so that in and of itself it describes one and only one feature of the plant. Third, there must be a way to repeat character-states or modifiers without repeating the character or character-state, respectively. This is especially important if variations in the statement of a character or a character-state are to be avoided when multiple values are appended. Fourth, the processing strategy must provide a means of compressing the data. Otherwise, the size of the data base would quickly get out of hand as the bulky descriptions, such as the one illustrated in Table 1, were compiled for upwards of 20,000 species, and the cost of processing would become astronomical.

The solution adopted by FNA is to generate a parallel numerical classification or numericlature and then to store and process the character statements as numerical codes using software capable of supporting hierarchical data structures. Only the actual character-states and their modifiers are stored as words in the individual records. The numerical codes constitute the cross-links to the Morphological and Ecological Vocabulary (MEV), the authority file governing the classification and order of the characters, where the numericlature is equated to its parallel nomenclature.

The numericlature is generated by assigning a unique, systematic, numerical code of fixed length to every unique term in each of the four levels of the character hierarchy, i.e., primary part, secondary part, tertiary part, and character (e.g., BRANCHES, INTERNODES, VESTITURE, VESTITURE TYPE). For each level or column (see Table 1), the vocabulary of terms is established in some absolutely fixed order, and then the terms are numbered consecutively from beginning to end. The numbering is done by tens or hundreds to allow for the interpolation of new terms in the future as needed without changing the relative order of the original terms. Thus, assuming for the sake of illustration that the terms on the first page of the data form shown in Figure 1 represent a closed vocabulary and that the numbering is done by tens, we have five terms in Column (Level) 1 that code as follows:

PLANTS	10
LEAVES	20
BASAL LEAVES	30
CAULINE LEAVES	40
SUBMERGED LEAVES	50

Any term interpolated between PLANTS and LEAVES will have a number between 10 and 20, say 15, etc.

In theory, at least, the original numbers are not changed once their assignment is definitive, i.e., after all editorial vetting and debugging, regardless of how many additions or deletions. If terms are deleted, their equivalent codes are deleted, and the remaining terms and codes retain their original relative order. This can only work if in the original numbering rather large gaps are left in the codes between successive terms, to allow for all possible future additions, and if the final order can be established at once. In practice, the ordering becomes the most difficult problem, and it may be necessary to regenerate the numericlature several times over a period of time before the order is satisfactory. Until data banking based on the numericlature has begun, regenerating the codes does not pose an

extraordinary problem. Once data banking has begun, however, then making a change in the code of a particular term in the authority file (i.e., the MEV) necessitaties changing the code wherever this term has been used in other files in the data bank. The impact of such changes is at once apparent. It multiplies rapidly as the files grow.

Returning to our example in Figure 1, there are three unique terms in Column (Level) 2, which code as follows:

BLADES	10
LEAFLETS	20
PETIOLES	30

The two unique terms in Column 3 and their codes are:

ADAXIAL SURFACES	10
ABAXIAL SURFACES	20

The fourth level, the Character Level, presents the greatest difficulty. The number of unique terms is much larger, although more limited than might be supposed, but the ordering is highly problematic. To illustrate without arguing order, let us take the order given in Figure 1.

DURATION	010
HABIT	020
HEIGHT	030
SEX	040
PRESENCE	050
LEAF TYPE	060
ARRANGEMENT	070
NUMBER	080
OUTLINE	090
LENGTH	100
WIDTH	110
APEX SHAPE	120
BASE SHAPE	130
MARGIN TYPE	140
VESTITURE TYPE	150

The full numericlature of a character statement consists of the set of codes for each of the four levels of hierarchy taken together--the numbers for a row of terms as presented in Table 1. Examples based on the above codes generated for Figure 1 are:

10 00 00 010	(DURATION of PLANTS)
20 10 00 090	(OUTLINE of BLADES of LEAVES)
40 10 00 090	(OUTLINE of BLADES of CAULINE LEAVES)
40 20 20 150	(VESTITURE TYPE of ABAXIAL SURFACES of CAULINE LEAVES)
40 30 00 100	(LENGTH of PETIOLES of CAULINE LEAVES)

Given this approach to the coding, which assigns a unique number from a consecutive sequence to each unique term at each level, the numericlature guarantees that the unique terms of a particular column (level) will always sort numerically into the same order. Likewise, it guarantees that a single order will always result when the terms are sorted hierarchically

by number, i.e., second-level categories (terms) within first-level categories, third-level within second-level categories, and fourth-level within third-level. No matter how many times a term (category), e.g., OUTLINE, repeats itself among categories of a higher level, it will always sort into the same order relative to other terms. To illustrate, the terms NUMBER, OUTLINE, and MARGIN TYPE of the fourth-level (character-level) vocabulary may be used repeatedly to describe different parts of the plant as delineated in the categories of the third or higher levels, but always NUMBER will precede OUTLINE and both will precede MARGIN TYPE within category. Furthermore, the relative order of terms in any subset, regardless how large or small, will be identical to that of the entire vocabulary, whether the columns are taken individually or collectively. Finally, within level a particular numerical code can have only one nomenclatural equivalent, whereas across two or more levels a particular combination of codes, i.e., a particular numericlatural value, can point to only one combination of terms. Thus, the codes can serve in place of the words for the purposes of search and comparison.

In the data bank, the characters of a plant are designated by their appropriate numerical codes, and the actual values or character-states and their modifiers are appended , as words either in full or abbreviated. Some examples based on Figure 1 and the above codes might be:

```
10 00 00 010   PERENNIAL
                 USUALLY
               BIENNIAL
                 SOMETIMES
    (DURATION of PLANTS USUALLY PERENNIAL, SOMETIMES BIENNIAL)

20 10 00 090   OVATE
                 or
               LANCEOLATE
                 BROADLY
    (OUTLINE of BLADES of LEAVES OVATE or BROADLY LANCEOLATE)

40 20 20 150   PUBERULENT
                 FINELY
    (VESTITURE TYPE of ABAXIAL SURFACES of CAULINE LEAVES FINELY
       PUBERULENT)
```

The numerical codes are further augmented in the computer record by coded printing instructions, as needed. To the code in each column must be added one more digit or a letter to indicate whether the term represented by the code (e.g., 10 = PLANTS) should be printed automatically by machine when a description is composed. Further coding may be necessary in the fourth column or following the character-state or its modifier to indicate what type of punctuation or connectives are needed and whether the modifier should precede or follow the character-state. Thus, in the example above concerning plant duration, the machine must be able to transform 10 00 00 010 PERENNIAL USUALLY BIENNIAL SOMETIMES into PLANTS USUALLY PERENNIAL, SOMETIMES BIENNIAL. This involves printing the first-level term, PLANTS, and the character-states and their modifiers properly arranged and punctuated, but not printing the second-, third-, and fourth-level terms, which in the first two instances are blanks but in the third case would be the word DURATION. The print codes will not be constant for any partic-

ular terms but presumably will be constant for particular combination of terms. How the print procedure works can be appreciated by studying the italicized words in Table 1.

The FNA character classification (Porter et al. 1973) consists of approximately 28,000 lines or rows of unique combinations of first-, second-, third-, and fourth-level terms. In order to allow for all of the potentially numerous new terms that might be added to each level, the original terms were numbered consecutively by hundreds, necessitating a five-digit code at each level. Thus the basic numeriClature consists of 20 digits, and following are some examples:

SECOND GLUME	AWNS	--	LENGTH	04770 02350 00000 07700
SECOND GLUME	AWNS	TRICHOMES	LENGTH	04770 02350 01400 07700
CALYX	--	--	POSITION	05004 00000 00000 05000
CALYX	LOBES	MARGIN TRICHOMES	LENGTH	05004 01800 04000 07700
CALYX	APPENDAGES	MARGIN TRICHOMES	SHAPE	05004 02900 04000 09100

It should be added that these codes are intended for internal use in the data-banking operation. The outside contributor or user need not concern himself with the codes, and for this reason they are not printed in Part II of the Guidebook along with the classified vocabulary.

Figure 2 is a schematic representation of the logical data hierarchy (inverted "branching tree") from the point of view of the processing system, GIS. The hierarchical file structures supported by GIS are explained in detail by Krauss (1973a, 1973b). It should be emphasized that the data hierarchy in this instance is not the same as the character hierarchy or nesting, nor does it have anything to do with taxonomic hierarchy.

This is a three-level hierarchy. The Fixed or Root Segment specifies the plant character being described, according to the four-level classification of characters explained above. In the authority file, i.e., the MEV, the terms and the equivalent numeriClature are given; in the Taxon Data Bank only the numeriClature is given. Segment 1_i represents the second level of the data hierarchy. This segment carries the term(s) for describing a particular character-state. It may repeat relative to the Root Segment as often as necessary to describe all the character-states (Segment 1_1, 1_2, 1_3, ... 1_n) of the specified character. Segment 2_i represents the third and lowest level of the data hierarchy which carries the modifiers. It may repeat relative to a particular character-state (e.g., Segment 1_1) as often as necessary (Segment 2_1, 2_2, 2_3, ... 2_n) to modify the given character-state. Herein lies the branching tree: one character may be described by many character-states each of which in turn may be qualified by many modifiers.

REFERENCES CITED

Krauss, H. M.

1973a The use of generalized information processing systems in the biological sciences. Taxon 22: 3-18. (Flora North America Report No. 67).

1973b The information system design for the Flora North America Program. Brittonia 25: 119-134. (Flora North America Report No. 69).

Porter, D. M., R. W. Kiger, and J. E. Monahan

1973 A guide for contributors to Flora North America, Part II: An outline and glossary of terms for morphological and habitat description (provisional edition). Flora North America Report No. 66, p. i-x, 1-120, G1-G32.

Shetler, S. G., et al.

1973 A guide for contributors to Flora North America (FNA) (provisional edition). Flora North America Report No. 65, p. i-ix, 1-28 + appendices A-E.

Figure 1. Standard FNA form for compiling morphological data on a taxon-by-taxon basis (first of four pages).

FLORA NORTH AMERICA

Family

TDB No.

Editor

Morphological Data Form

Genus Species Subsp./Var. Contributor Date

Sequence and terminology of plant parts, characters and character states should follow the Morphological Outline and Glossary in FNA Guidebook Part II. Complete one form for every taxon, regardless of rank.

DIAGNOSTIC DESCRIPTION

COMPARATIVE DESCRIPTION (Minimum Data Requirements)

PLANTS
- duration
- habit
- height
- sex

LEAVES
- presence
- leaf type
- arrangement

BLADES
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

LEAFLETS
- number
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

PETIOLES
- presence
- length
- vestiture type

BASAL LEAVES
- presence
- leaf type
- arrangement

BLADES
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

LEAFLETS
- number
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

PETIOLES
- presence
- length
- vestiture type

CAULINE LEAVES
- presence
- leaf type
- arrangement

BLADES
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

LEAFLETS
- number
- outline
- length
- width
- apex shape
- base shape
- margin type

Adaxial surfaces
- vestiture type

Abaxial surfaces
- vestiture type

PETIOLES
- presence
- length
- vestiture type

SUBMERGED LEAVES
- presence
- leaf type
- arrangement

BLADES
outline
length
width
apex shape
base shape
margin type
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type
LEAFLETS
number
outline
length
width
apex shape
base shape
margin type
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type
PETIOLES
presence
length
vestiture type
FLOATING LEAVES
presence
leaf type
arrangement
BLADES
outline
length
width
apex shape
base shape
margin type
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type
LEAFLETS
number
outline
length
width
apex shape
base shape
margin type
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type
PETIOLES
presence
length
vestiture type
EMERGENT LEAVES
presence
leaf type
arrangement
BLADES
outline
length
width
apex shape
base shape
margin type
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type

LEAFLETS
number
outline
length
width
apex shape
base shape
margin shape
Adaxial surfaces
vestiture type
Abaxial surfaces
vestiture type
PETIOLES
presence
length
vestiture type
STIPULES
presence
INFLORESCENCES
position
primary arrangement
secondary arrangement
MIXED INFLORESCENCES
position
primary arrangement
secondary arrangement
STAMINATE INFLORESCENCES
position
primary arrangement
secondary arrangement
CARPELLATE INFLORESCENCES
position
primary arrangement
secondary arrangement
FLOWERS
PERIANTH or CALYX
presence
fusion
symmetry
length
Lobes
number
outline
color
length
Outer vestiture
vestiture type
TEPALS or SEPALS
presence
number
fusion
symmetry
color
length
Abaxial surfaces
vestiture type
COROLLA
presence
fusion
symmetry
color
length
Lobes
number
outline
color
length
Outer vestiture
vestiture type
PETALS
presence
number
fusion
outline

PETALS (con't.)
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

STAMENS
- number
- number of whorls
- insertion
- arrangement

FILAMENTS
- fusion

GYNOECIUM

Ovaries
- number
- position

Carpels
- number
- fusion

Locules
- number

Placentae
- position

Styles
- number

Style branches
- number

BISEXUAL FLOWERS

PERIANTH or CALYX
- presence
- fusion
- symmetry
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

TEPALS or SEPALS
- presence
- number
- fusion
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

COROLLA
- presence
- fusion
- symmetry
- color
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

PETALS
- presence
- number
- fusion
- outline
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

STAMENS
- number
- number of whorls
- insertion
- arrangement

FILAMENTS
- fusion

GYNOECIUM

Ovaries
- number
- position

Carpels
- number
- fusion

Locules
- number

Placentae
- position

Styles
- number

Style branches
- number

STAMINATE FLOWERS or DISC FLORETS

PERIANTH or CALYX
- presence
- fusion
- symmetry
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

TEPALS or SEPALS
- presence
- number
- fusion
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

COROLLA
- presence
- fusion
- symmetry
- color
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

PETALS
- presence
- number
- fusion
- outline
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

STAMENS
- number
- number of whorls
- insertion
- arrangement

FILAMENTS
- fusion

CARPELLATE FLOWERS or RAY FLORETS

PERIANTH or CALYX
- presence
- fusion
- symmetry
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

TEPALS or SEPALS
- presence
- number
- fusion
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

COROLLA
- presence
- fusion
- symmetry
- color
- length

Lobes
- number
- outline
- color
- length

Outer vestiture
- vestiture type

PETALS
- presence
- number
- fusion
- outline
- symmetry
- color
- length

Abaxial surfaces
- vestiture type

GYNOECIUM

Ovaries
- number
- position

Carpels
- number
- fusion

Locules
- number

Placentae
- position

Styles
- number

Style branches
- number

FRUITS
- fruit type
- length
- width
- dehiscence
- vestiture type

SEEDS
- number per fruit
- length
- width

CHROMOSOME NUMBER (optional)

ENTER chromosome counts only as somatic (2n) numbers. Counts must be based on material of wild origin from FNA territory. Document each number with one reliable primary literature report. Cite title of journal or monograph, volume and page, and year. Continue listing under "Additional Data" or on separate sheet, if necessary. Do not include unpublished counts.

No. Citation: title, volume & page, year

No. Citation: title, volume & page, year

No. Citation: title, volume & page, year

ADDITIONAL DATA (optional)

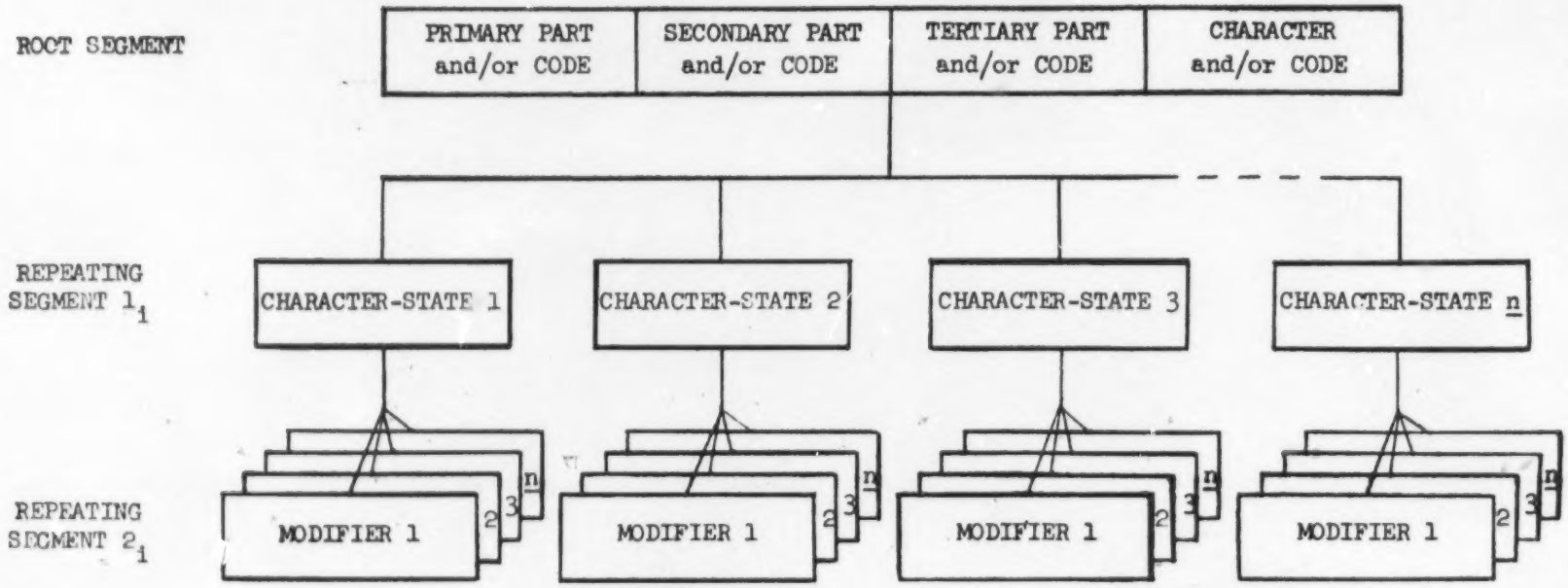

Figure 2. Schematic representation of the logical, three-level, data hierarchy for a morphological description in the Taxon Data Bank (TDB) as structured for processing by the Generalized Information System (GIS). The Root Segment alone is used in the backup authority file, the Morphological and Ecological Vocabulary (MEV), but with both words and codes, whereas all levels are used in the TDB, but only codes are entered in the Root Segment.

Table 1. A partial description of Xylosma flexuosa (H.B.K.) Hemsl. (Flacourtiaceae) in FNA Data Bank format (see Guidebook, Part II, Porter et al. 1973)

Primary Part	Secondary Part	Tertiary Part	Character	Character-State	Modifier	Guidebook Reference
PLANTS	--	--	DURATION	PERENNIAL	--	p. 1
PLANTS	--	--	HABIT	*SHRUBS**[1]	--	p. 1
PLANTS	--	--	HEIGHT	*2 M[HIGH]**[2]	*TO**[3][UP TO]	p. 1
PLANTS	--	--	SEX	*POLYGAMODIOECIOUS**	--	p. 1
MAIN STEMS	*WOOD*	--	TEXTURE	*HARD*	*MODERATELY*[3]	p. 1
MAIN STEMS	WOOD	--	WOOD POROSITY	*DIFFUSE-POROUS*	--	p. 1
BRANCHES	--	--	PRESENCE	PRESENT	--	p. 2
BRANCHES	NODES	--	DIAMETER	*SLENDER*	[*OR*]	p. 2
				STOUT		
BRANCHES	NODES	VESTITURE	VESTITURE TYPE	*GLABRATE*	--	p. 2
BRANCHES	INTERNODES	--	DIAMETER	SLENDER	[OR]	p. 2
				STOUT		
BRANCHES	INTERNODES	VESTITURE	VESTITURE TYPE	GLABRATE	--	p. 2
TWIGS	--	--	PRESENCE	PRESENT	--	p. 3
TWIGS	NODES	VESTITURE	VESTITURE TYPE	*PUBERULENT*	*DENSELY*[3][*OR*]	p. 3
				GLABRESCENT		
TWIGS	INTERNODES	VESTITURE	VESTITURE TYPE	PUBERULENT	DENSELY[OR]	p. 3
				GLABRESCENT		
SPURS	--	--	PRESENCE	*PRESENT*	*OFTEN*[3]	p. 10
*THORNS**	--	--	PRESENCE	*PRESENT**	*OFTEN**[3]	p. 12
THORNS	--	--	LENGTH	*33 MM[LONG]*	*TO*[3][UP TO]	p. 12
THORNS	NODES	--	DIAMETER	*SLENDER*	[*OR*]	p. 12
				STOUT		
THORNS	INTERNODES	--	DIAMETER	SLENDER	[OR]	p. 12
				STOUT		

Primary Part	Secondary Part	Tertiary Part	Character	Character-State	Modifier	Guidebook Reference
LEAVES	--	--	PRESENCE	PRESENT	--	p. 14
LEAVES	--	--	PROXIMITY	*REMOTE*	[*OR*]	p. 14
LEAVES	--	--	ARRANGEMENT	*CROWDED (ON SPURS)*	*SOMETIMES*[3]	p. 14
BLADES	--	--	OUTLINE	*RHOMBIC-OVATE**	[,]*	p. 14
				*OVATE-ELLIPTIC**	[,]	p. 14
				*ELLIPTIC**	[,]	p. 14
				*OBLANCEOLATE**	[*OR*]	p. 14
				*OBOVATE**	--	p. 14
BLADES	--	--	TEXTURE	*CHARTACEOUS**	[*OR*]*	p. 14
				*SUBCORIACEOUS**		
BLADES	--	--	LENGTH	*6 CM*[*LONG*]*	*TO**[3][UP TO]	p. 14
BLADES	--	--	WIDTH	*3 CM*[*WIDE*]*	*TO**[3][UP TO]	p. 14
BLADES	*APEX**	--	APEX SHAPE	*ACUMINATE**	[,]*	p. 14
				*ACUTE**	[,]*	
				*OBTUSE**	[*OR*]*	
				*ROUNDED**	--	
BLADES	*BASE**	--	BASE SHAPE	*ATTENUATE-CUNEATE**	[*OR*]*	p. 14
				CUNEATE		
BLADES	*MARGIN**	--	MARGIN TYPE	*CRENULATE-SERRULATE**	*IRREGULARLY**[3]	p. 14
BLADES	*ADAXIAL SURFACES*	SURFACE	COLOR	*GREEN*	--	p. 15
BLADES	ADAXIAL SURFACES	SURFACE	SURFACE CHARACTER	*LUSTROUS*	*OFTEN*[3]	p. 15
BLADES	ADAXIAL SURFACES	VESTITURE	VESTITURE TYPE	*GLABROUS*	--	p. 15
BLADES	ADAXIAL SURFACES	*PRIMARY VEINS*	VESTITURE TYPE	*PUBERULENT*	--	p. 15
BLADES	ADAXIAL SURFACES	*SECONDARY VEINS*	GENERAL	*EVIDENT*	--	p. 15

Primary Part	Secondary Part	Tertiary Part	Character	Character-State	Modifier	Guidebook Reference
BLADES	ADAXIAL SURFACES	SECONDARY VEINS	NUMBER PER SIDE	*5-6[PER SIDE]*	--	p. 15
BLADES	ADAXIAL SURFACES	*TERTIARY VEINS*	GENERAL	*INCONSPICUOUS*	--	p. 15
BLADES	*ABAXIAL SURFACES*	SURFACE	COLOR	*GREEN*	*PALER*[3]	p. 15
BLADES	ABAXIAL SURFACES	VESTITURE	VESTITURE TYPE	*GLABROUS*	--	p. 15
BLADES	ABAXIAL SURFACES	*SECONDARY VEINS*	GENERAL	*EVIDENT*	--	p. 15
BLADES	ABAXIAL SURFACES	*TERTIARY VEINS*	GENERAL	*INCONSPICUOUS*	--	p. 15
*PETIOLES**	--	--	PRESENCE	PRESENT	--	p. 24
PETIOLES	--	--	LENGTH	*SHORT**	--	p. 24
*INFLORESCENCES**	--	--	INFL ARRANGEMENT	*FASCICULATE**	--	p. 29
INFLORESCENCES	--	--	POSITION	*SESSILE**	--	p. 29
*PEDICELS**	--	--	LENGTH	*5 MM[LONG]*	*TO**[3][UP TO]	p. 32
PEDICELS	--	VESTITURE	VESTITURE TYPE	*GLABRATE*	[*OR*]	p. 32
				PUBERULENT		
*SEPALS**	--	--	PRESENCE	PRESENT	--	p. 65
SEPALS	--	--	NUMBER	*5**	--	p. 65
SEPALS	--	--	OUTLINE	*OVATE-OBLONG*	--	p. 65
				*1-2 MM[LONG]**	--	p. 65
SEPALS	*APEX*	--	APEX SHAPE	*ACUTE*	[*OR*]	p. 65
				OBTUSE	--	
SEPALS	*MARGIN**	--	MARGIN TYPE	*CILIATE**	*CONSPICUOUSLY**	p. 65

Primary Part	Secondary Part	Tertiary Part	Character	Character-State	Modifier	Guidebook Reference
SEPALS	*ABAXIAL SURFACES*	VESTITURE	VESTITURE TYPE	*GLABROUS*	[*OR*]	p. 65
				PUBERULENT	--	
*FILAMENTS**	--	--	LENGTH	*3-4 MM*[*LONG*]*	--	p. 68
GYNOECIUM	OVARIES	*CARPELS*	NUMBER	*2-6*	--	p. 71
GYNOECIUM	OVULES	--	*OVULE NUMBER PER LOCULE*	*2-8*	--	p. 71
GYNOECIUM	*STYLES*	--	DURATION	*PERSISTENT*	--	p. 71
GYNOECIUM	*STYLE BRANCHES**		NUMBER	*2**	--	p. 71
GYNOECIUM	*STIGMAS*	--	NUMBER	*2*	--	p. 71
GYNOECIUM	STIGMAS	--	SHAPE	*DILATED*	--	p. 71
*FRUITS**	--	--	SHAPE	*SUBGLOBOSE**	--	p. 118
FRUITS	--	--	DIAMETER	*6 MM*[*DIAMETER*]*	*TO**[3][UP TO]	p. 118
FRUITS	--	SURFACE	COLOR	*RED**	--	p. 118
FRUITS	--	VESTITURE	VESTITURE TYPE	*GLABROUS**	--	p. 118
*SEEDS**	--	--	LENGTH	*3-4 MM*[*LONG*]	--	p. 119
SEEDS	--	--	DIAMETER	*2-3 MM*[*DIAMETER*]	--	p. 119
SEEDS	--	--	*SEED NUMBER PER FRUIT*	*2-8**	--	p. 119

[1]The terms, modifiers, and connectives (including punctuation) printed in *italics* constitute the description as it would be presented in written form (reading from left to right and from top to bottom). These are the words and marks that would be coded for print out. The diagnostic description consists of the subset of these terms marked by asterisks (*).
[2]Interpolated words are placed in brackets [] and italicized where essential to the sense of the descriptive statement.
[3]These modifiers should be placed before rather than after the character-state when printed.